

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 17:14:36 ; Search time 78.5143 Seconds  
(Without alignments)  
504.414 Million cell updates/sec

Title: US-09-807-933b-3  
Perfect score: 2020  
Sequence: 1 MKRITITSSALLALALGTEM.....TYKEVTPCKEITAKTGSRK 366

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PTCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	757.5	37.5	229	9	US-10-007-521-12
2	753.5	37.3	225	9	US-10-007-521-2
3	753.5	37.3	297	9	US-10-007-521-4
4	753.5	37.3	308	9	US-10-007-521-6
5	748.5	37.1	205	9	US-09-261-329-7
6	742.5	36.8	201	9	US-09-261-329-5
7	741.5	36.7	201	9	US-09-261-329-4
8	722.5	35.8	203	9	US-09-261-329-9
9	722.5	35.8	222	9	US-10-007-521-14
10	722.5	35.8	294	9	US-10-007-521-24
11	715	35.4	349	9	US-10-007-521-10
12	706	35.0	310	9	US-10-007-521-22
13	698	34.6	202	9	US-09-261-329-3
14	690.5	34.2	235	1	US-08-841-636A-31
15	683.5	33.8	376	10	US-09-735-787-4
16	675.5	33.4	203	9	US-09-261-329-6
17	675	33.4	305	10	US-09-735-787-2
18	674	33.4	202	9	US-09-261-329-1
19	665.5	32.9	226	9	US-10-007-521-16

20	665.5	32.9	293	9	US-10-007-521-20	Sequence 20, Appl
21	665.5	32.9	298	9	US-10-007-521-18	Sequence 18, Appl
22	662	32.8	295	9	US-10-007-521-8	Sequence 8, Appl
23	660.5	32.7	203	9	US-09-261-329-8	Sequence 8, Appl
24	654	32.4	202	9	US-09-261-329-2	Sequence 2, Appl
25	506.5	25.1	211	9	US-09-261-329-11	Sequence 11, Appl
26	492.5	24.4	235	9	US-09-261-329-1	Sequence 10, Appl
27	440.5	21.8	138	9	US-10-007-521-26	Sequence 26, Appl
28	220.5	10.9	75	9	US-10-007-521-32	Sequence 32, Appl
29	216.5	10.7	493	9	US-10-007-294A-2	Sequence 2, Appl
30	190.5	9.4	57	9	US-10-007-521-52	Sequence 52, Appl
31	186.5	9.2	59	9	US-10-007-521-34	Sequence 34, Appl
32	180.5	8.9	60	9	US-10-007-521-42	Sequence 42, Appl
33	175.5	8.7	60	9	US-10-007-521-70	Sequence 70, Appl
34	174	8.6	39	9	US-09-916-494A-4	Sequence 4, Appl
35	173.5	8.6	3732	9	US-10-123-155-71	Sequence 71, Appl
36	172.5	8.5	1591	9	US-10-073-912-13	Sequence 13, Appl
37	172	8.5	773	9	US-10-184-634-429	Sequence 429, App
38	172	8.5	773	9	US-10-184-634-429	Sequence 429, App
39	171	8.5	1422	10	US-09-735-933-1	Sequence 1, Appl
40	171	8.5	2403	9	US-10-184-644-45	Sequence 45, Appl
41	171	8.5	2403	9	US-10-184-634-45	Sequence 45, Appl
42	169	8.4	4372	9	US-10-123-155-53	Sequence 53, Appl
43	168	8.3	18636	9	US-10-073-912-17	Sequence 17, Appl
44	167.5	8.3	2207	9	US-10-123-155-137	Sequence 137, App
45	166.5	8.2	1730	9	US-10-123-155-7	Sequence 7, Appl

#### ALIGNMENTS

RESULT 1  
US-10-007-521-12  
Sequence 12, Application US/10007521  
Publication No. US20030054539A1  
GENERAL INFORMATION:  
APPLICANT: Schuilein, Martin  
Laessen, Soren F.  
Kauppinen, Markus S.  
Lange, Lene  
Nielsen, Ruby I.  
Ihara, Michiko  
Takagi, Shinobu  
TITLE OF INVENTION: No. US20030054539A1 Endoglucanases  
NUMBER OF SEQUENCES: 109  
CORRESPONDENCE ADDRESS:  
ADDRESSER: No. US20030054539A10 No. US20030054539A1disk of No. US20030054539A1  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/007,521  
FILING DATE: 10-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/651,136  
FILING DATE: 21-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4366,200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:  
LENGTH: 299 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-007-521-12

Query Match 37.5%; Score 757.5; DB 9; Length 299;  
Best Local Similarity 62.3%; Pred. No. 4.8e-44;  
Matches 134; Conservative 32; Mismatches 40; Indels 9; Gaps 4;

QY 152 IVSGASGNGVTRVWDCCKASGWPGRKAVSSPVKSCNKGVTALSDSNVQGCN GNS 211  
DB 17 LVASASGSGSTRYWPCCKRSCAMPKAAVSOQVYACDAN-FQRLSDFNVOGSGNGGSA 75  
QY 212 YMCNDNPMVNDLAVGFAAAIISGGESRWCCSCPELFTSTSVAGKKMVIQVNTTGG 271  
DB 76 YSCADQTPWAVNDLAVGFAATSTAGSESSWCACALFTTSGPVAGKTMVQSTSTGG 135  
QY 272 DLGSGTAHFDLQMPGGGVCIFNGCSKQWGA-PNDGWSRYGGISASDCSLPSALQAG 330  
DB 136 DLGSGN---QPDIAHFGGVCIFNGCSSQFGSLP---GAQYGGISSRDQCDSPFAPLPG 188  
QY 331 CKRPFNMFKNADNPSMTYKEYTCPEKITAKTGCSR 365  
DB 189 COMRFDMFONADNPTFTFQVQCPAEIVARSGCSR 223

RESULT 2  
US-10-007-521-2

Sequence 2, Application US/10007521  
Publication No. US20030054539A1

## GENERAL INFORMATION:

APPLICANT: Schulein, Martin

Andersen, Lene N.

Laessen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takagi, Shinobu

TITLE OF INVENTION: No. US20030054539A1el Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESSER: No. US20030054539A1o No. US20030054539A1disk of No. US200300545

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,521

FILING DATE: 10-Dec-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366,200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 225 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-007-521-2

Query Match 37.3%; Score 753.5; DB 9; Length 225;  
Best Local Similarity 60.8%; Pred. No. 6.6e-44;  
Matches 127; Conservative 36; Mismatches 41; Indels 5; Gaps 3;

QY 158 SGNGVTRVWDCCKASGWPGRKAVSSPVKSCNKGVTALSDSNVQGCN GNSYMCND 216  
DB 21 SGIGQTRVWDCCKRSCAMPK-GPSSPVQACDKNDNPLNDGSGSTRSCDAGSAYMCS 79  
QY 217 NOPAVNDLAVGFAAAIISGGESRWCCSCPELFTSTSVAGKKMVIQVNTTGGDLSS 276  
DB 80 QSPWAVSELSYGMVAVLVLAGSSQWCACCELTFTSGPVAGKKMVIQVNTTGGDLSDN 139  
QY 277 TGAHFDLQMPGGGVCIFNGCSKQWGA-PNDGWSRYGGISASDCSLPSALQAGCKMRFN 336  
DB 140 ---HFDLAIPEGGVCIFNACTQYGAFPNGMDRYGIIHSKEGCSPEALKKPGCMRFD 196  
QY 337 WFNADNPSMTYKEYTCPEKITAKTGCSR 365  
DB 197 WFNADNPSVTFQEVACPELTSKSGCSR 225

RESULT 3  
US-10-007-521-4

Sequence 4, Application US/10007521  
Publication No. US20030054539A1

## GENERAL INFORMATION:

APPLICANT: Schulein, Martin

Andersen, Lene N.

Laessen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takagi, Shinobu

TITLE OF INVENTION: No. US20030054539A1el Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESSER: No. US20030054539A1o No. US20030054539A1disk of No. US200300545

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,521

FILING DATE: 10-Dec-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366,200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 297 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-007-521-4

Query Match 37.3%; Score 753.5; DB 9; Length 297;  
Best Local Similarity 60.8%; Pred. No. 8.9e-44;  
Matches 127; Conservative 36; Mismatches 41; Indels 5; Gaps 3;

QY 158 SGNVTRTWDCCKKASCSMPKGNVSSPVKSCNKGVTALSDSNVQSGN-GENSYMCDND 216  
DB 21 SGIGGTRTWDCCKKSCAMPK-GPSSPVQACDKNDPLNDGSTRSGCDAGSAYMCS 79  
QY 217 NOPMAYNDNLAYGFAAAAISSGGSRWCCSCFELFTSTSVAGKXVIQVNTGDLGSS 276  
DB 80 SPMAVSDLSYGMVAAYKLAGSSESQMCACVETLTSSPVAGKXVIQVNTGDLGDN 139  
QY 277 TGAHFDLQMPGGVGIFNGCSKQMGAPNDGMSRYGSISSASDCSLPSALQAGCKWRPN 336  
DB 140 ---HFDLAIPGGVGIFNACTDQYGAPPMGMDRYGSIHSKECESFPALKPGCNMFRD 196  
QY 337 WPKNADNPMTYKEVTCPEKITAKTGCSR 365  
DB 197 WFOADNPSTVTFQEVACPSBELTSKSGCSR 225

RESULT 4  
US-10-007-521-6

Sequence 6, Application US/10007521  
Publication No. US20030054539A1  
GENERAL INFORMATION:

APPLICANT: Schuelein, Martin

Andersen, Lene N.

Laessen, Soren P.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Thara, Michiko

Takagi, Shinobu

TITLE OF INVENTION: No. US20030054539A1e1 Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESS: No. US20030054539A1o No. US20030054539A1disk of No. US200300545

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/10/007,521

APPLICATION DATE: 10-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,136

FILING DATE: 21-May-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambirth, Elias J.

REGISTRATION NUMBER: 33,728

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 308 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-007-521-6

Query Match 37.3%; Score 753.5; DB 9; Length 308;  
Best Local Similarity 60.8%; Pred. No. 9.2e-44;  
Matches 127; Conservative 36; Mismatches 41; Indels 5; Gaps 3;

QY 158 SGNVTRTWDCCKKASCSMPKGNVSSPVKSCNKGVTALSDSNVQSGN-GENSYMCDND 216  
DB 21 SGIGGTRTWDCCKKSCAMPK-GPSSPVQACDKNDPLNDGSTRSGCDAGSAYMCS 79  
QY 217 NOPMAYNDNLAYGFAAAAISSGGSRWCCSCFELFTSTSVAGKXVIQVNTGDLGSS 276  
DB 80 SPMAVSDLSYGMVAAYKLAGSSESQMCACVETLTSSPVAGKXVIQVNTGDLGDN 139  
QY 277 TGAHFDLQMPGGVGIFNGCSKQMGAPNDGMSRYGSISSASDCSLPSALQAGCKWRPN 336  
DB 140 ---HFDLAIPGGVGIFNACTDQYGAPPMGMDRYGSIHSKECESFPALKPGCNMFRD 196  
QY 337 WPKNADNPMTYKEVTCPEKITAKTGCSR 365  
DB 197 WFOADNPSTVTFQEVACPSBELTSKSGCSR 225

RESULT 5  
US-09-261-329-7

Sequence 7, Application US/09261329  
Publication No. US20030092097A1  
GENERAL INFORMATION:

APPLICANT: Andersen, Kim

APPLICANT: Schuelein, Martin

APPLICANT: Christiansen, Lars

APPLICANT: Damgaard, Bo

APPLICANT: von der Osten, Claus

TITLE OF INVENTION: Cellulase Variants

FILE REFERENCE: 4887.204-US

CURRENT APPLICATION NUMBER: US/09/261,329

CURRENT FILING DATE: 1999-03-03

EARLIER APPLICATION NUMBER: 1013/96

NUMBER OF SEQ ID NOS: 26

SOFTWARE: FaastSeq for Windows Version 3.0

SEQ ID NO 7

LENGTH: 205

TYPE: PRT

ORGANISM: Cellulase variants

US-09-261-329-7

Query Match 37.1%; Score 748.5; DB 9; Length 205;  
Best Local Similarity 60.6%; Pred. No. 1.3e-43;  
Matches 126; Conservative 36; Mismatches 41; Indels 5; Gaps 3;

QY 159 SGNVTRTWDCCKKASCSMPKGNVSSPVKSCNKGVTALSDSNVQSGN-GENSYMCDND 217  
DB 1 SGIGGTRTWDCCKKSCAMPK-GPSSPVQACDKNDPLNDGSTRSGCDAGSAYMCSQ 59  
QY 218 QPMAYNDNLAYGFAAAAISSGGSRWCCSCFELFTSTSVAGKXVIQVNTGDLGSS 277  
DB 60 SPMAVSDLSYGMVAAYKLAGSSESQMCACVETLTSSPVAGKXVIQVNTGDLGDN- 118  
QY 278 GAHFDLQMPGGVGIFNGCSKQMGAPNDGMSRYGSISSASDCSLPSALQAGCKWRPN 337  
DB 119 --HFDLAIPGGVGIFNACTDQYGAPPMGMDRYGSIHSKECESFPALKPGCNMFRD 176  
QY 338 FPNADNPMTYKEVTCPEKITAKTGCSR 365  
DB 177 FPNADNPSTVTFQEVACPSBELTSKSGCSR 204

RESULT 6  
US-09-261-329-5

Sequence 5, Application US/09261329  
Publication No. US20030092097A1  
GENERAL INFORMATION:

Query Match	36.8%;	Score	742.5;	DB	9;	Length	201;
Best Local Similarity	63.0%;	Pred. No.	3.3e-43;				
Matches	131;	Conservative	30;	Mismatches	38;	Indels	9;
						Gaps	4.

RESULT 7  
US-09-261-329-4  
; Sequence 4, Application US/09261329  
; Publication No. US20030092097A1

Query Match	35.7%	Score 741.5;	DB 9;	Length 201;
Best Local Similarity	60.6%;	Pred. No. 3.8e+3;		
Matches 126; Conservative	40;	Mismatches 33;	Indels 9;	Gaps 4

219 PWA VNDNLAYGF AAAAAISGGGESRWCSCFELTFTSTSVAGKKNVIQVNTNGDLSSTG 270

RESULT 8  
US-09-261-329-9

Query Match	35.8%;	Score 722.5;	DB 9;	Length 203;
Best Local Similarity	59.4%;	Pred. No. 7.4e-42;		
Matches 123;	Conservative 41;	Mismatches 36;	Indels 7;	Gaps

RESULT 9  
US-10-007-521-14  
; Sequence 14, Application US/10007521  
; Publication No. US20030054539A1

TITLE OF INVENTION: No. US20030054539A1 Endoglycanases  
 NUMBER OF SEQUENCES: 109  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: No. US20030054539A10 No. US20030054539A1disk of No. US20030054539A1  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/007,521  
FILING DATE: 10-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/651,136  
FILING DATE: 21-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4366.200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-10-007-521-14

Query Match 35.8%; Score 722.5; DB 9; Length 222;  
Best Local Similarity 59.4%; Pred. No. 8.2e-42;  
Matches 123; Conservative 41; Mismatches 36; Indels 7; Gaps 4;  
QY 160 NGVTTYMDCCKRSCSMPGKANVSPVKSCH-KDGVTLASDSNVGSCNGSGNSYMCNDNQ 218  
DB 21 SGVTTYRMDCCKRSCSMPGKANVSPVKSCH-KDGVTLASDSNVGSCNGSGNSYMCNDNQ 79  
QY 219 PAAVNDNLAYGPAALAIISGGESRWCCSCFELFTFTSTVAGKKMVIQVNTTGGDLGSSSTG 278  
DB 80 PAAVNDNLAYGPAALAIISGGESRWCCSCFELFTFTSTVAGKKMVIQVNTTGGDLGSSSTG 137  
QY 279 AHFDLMPGGGVCIFNGCSKQWCAIPNDGWSRYGSISSASDSSLPALQAGCKWRFNMF 338  
DB 138 -HFDIAMPGGGVCIFNGCSKQWCAIPNDGWSRYGSISSASDSSLPALQAGCKWRFNMF 194  
QY 339 KNAADNPSMTYKEVTCPEKREITAKTGCSR 365  
DB 195 ENADNPTVMEPVTCPELVAARTGCSR 221

RESULT 10  
US-10-007-521-24  
Sequence 24, Application US/10007521  
Publication No. US20030054539A1  
GENERAL INFORMATION:  
APPLICANT: Schultein, Martin  
Andersen, Lene N.  
Laessen, Soren F.  
Kauppinen, Markus S.  
Lange, Lene  
Nielsen, Ruby I.  
Ihara, Michiko  
Takagi, Shinobu  
TITLE OF INVENTION: No. US20030054539A1el Endoglucanases  
NUMBER OF SEQUENCES: 109  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. US20030054539A10 No. US20030054539A1disk of No. US20030054539A1  
STREET: 405 Lexington Avenue, 64th Floor

CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/007,521  
FILING DATE: 10-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/651,136  
FILING DATE: 21-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4366.200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 294 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-10-007-521-24

Query Match 35.8%; Score 722.5; DB 9; Length 294;  
Best Local Similarity 59.4%; Pred. No. 1.1e-41;  
Matches 123; Conservative 41; Mismatches 36; Indels 7; Gaps 4;  
QY 160 NGVTTYMDCCKRSCSMPGKANVSPVKSCH-KDGVTLASDSNVGSCNGSGNSYMCNDNQ 218  
DB 21 SGVTTYRMDCCKRSCSMPGKANVSPVKSCH-KDGVTLASDSNVGSCNGSGNSYMCNDNQ 79  
QY 219 PAAVNDNLAYGPAALAIISGGESRWCCSCFELFTFTSTVAGKKMVIQVNTTGGDLGSSSTG 278  
DB 80 PAAVNDNLAYGPAALAIISGGESRWCCSCFELFTFTSTVAGKKMVIQVNTTGGDLGSSSTG 137  
QY 279 AHFDLMPGGGVCIFNGCSKQWCAIPNDGWSRYGSISSASDSSLPALQAGCKWRFNMF 338  
DB 138 -HFDIAMPGGGVCIFNGCSKQWCAIPNDGWSRYGSISSASDSSLPALQAGCKWRFNMF 194  
QY 339 KNAADNPSMTYKEVTCPEKREITAKTGCSR 365  
DB 195 ENADNPTVMEPVTCPELVAARTGCSR 221

RESULT 11  
US-10-007-521-10  
Sequence 10, Application US/10007521  
Publication No. US20030054539A1  
GENERAL INFORMATION:  
APPLICANT: Schultein, Martin  
Andersen, Lene N.  
Laessen, Soren F.  
Kauppinen, Markus S.  
Lange, Lene  
Nielsen, Ruby I.  
Ihara, Michiko  
Takagi, Shinobu  
TITLE OF INVENTION: No. US20030054539A1el Endoglucanases  
NUMBER OF SEQUENCES: 109  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. US20030054539A10 No. US20030054539A1disk of No. US20030054539A1  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York

COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/007,521  
FILING DATE: 10-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/651,136  
FILING DATE: 21-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4366,200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 349 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-007-521-10  
Query Match 35.4%; Score 715; DB 9; Length 349;  
Best Local Similarity 58.3%; Pred. No. 4,2e-41;  
Matches 123; Conservative 36; Mismatches 42; Indels 10; Gaps 4;  
QY 157 ASGNVTRRYMDCCKASCSMPGKANVSSPVKSNKDGVTALSDSNVSGCNGNSYMCND 216  
DB 21 ASGKHTIRYMDCKTSCMBEGKASVSEVLTCKNQD-NPIVDNARSGCGGGAFACTN 79  
QY 217 NOPMAVNDNLAYGFAAAAIISGGESRWCCSCFELTFTSTVAGKQWVIOVNTTGGDLS 276  
DB 80 NSPMVASEDLAYGFAATATLSCGTESGWCACVATFTSGPVAQKMWVQSTNTGDDLSNN 139  
QY 277 TGAHFDLQMPGGGVGIFNGCSKQWQ--APNDGMSRYGSISSASDCSSLPALQAGCKR 334  
DB 140 ---HFLMTPGGGLGIFDGCSPAQFGLP---GERYGVSSRSQCDQMPELKDCCQWR 192  
QY 335 FNMFKADNPMSMTYKEVTCPEKITAKTGCSR 365  
DB 193 FDMFKSDNPDIIEFQVQCPKELIIVSGCVR 223  
RESULT 12  
US-10-007-521-22  
Sequence 22, Application US/10007521  
Publication No. US20030054539A1  
GENERAL INFORMATION:  
APPLICANT: Schuelein, Martin  
Andersen, Lene N.  
Larsen, Soren F.  
Kampinen, Markus S.  
Lange, Lene  
Nielsen, Ruby I.  
Thara, Michiko  
Takagi, Shinobu  
TITLE OF INVENTION: No. US20030054539A1 Endoglucanases  
NUMBER OF SEQUENCES: 109  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: No. US20030054539A10 No. US20030054539A1disk of No. US200300545  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/007,521  
FILING DATE: 10-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/651,136  
FILING DATE: 21-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4366,200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 310 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-10-007-521-22  
Query Match 35.0%; Score 706; DB 9; Length 310;  
Best Local Similarity 58.9%; Pred. No. 1.5e-40;  
Matches 126; Conservative 32; Mismatches 48; Indels 8; Gaps 5;  
QY 153 VSGASGNVTRRYMDCCKASCSMPGKANVSSPVKSNKDGVTALSDSNVSGCNGNS 211  
DB 14 VAQSSGTIRTRYMDCCKPSCGMEKASVSPVATCQRNN-NPIA-STASGCDNSGVA 71  
QY 212 YMCNDNOPMAVNDNLAYGFAAAAIISGGESRWCCSCFELTFTSTVAGKQWVIOVNTTGG 271  
DB 72 YTCNDNOPMAVNDNLAYGFAATAFBSGSEASWCACVATFTSGPVAQKMWVQSTNTG 131  
QY 272 DLAGSTGAHFDLQMPGGGVGIFNGCSKQWCAPNMGSRYGSISSASDCSSLPALQAGCK 331  
DB 132 DL---SGNFDLMPGGGLGIFDGCIPQWVFP--GNRYGTTSRSCSQIPALQPC 186  
QY 332 KRFNFKADNPMSMTYKEVTCPEKITAKTGCSR 365  
DB 187 NRYDWFNDADNDPVGSMRVRQCPALTDRTGCR 220  
RESULT 13  
US-09-261-329-3  
Sequence 3, Application US/09261329  
Publication No. US20030092097A1  
GENERAL INFORMATION:  
APPLICANT: Andersen, Kim  
Schuelein, Martin  
APPLICANT: Christiansen, Lars  
APPLICANT: Damgaard, Bo  
APPLICANT: Von Der Oesten, Claus  
TITLE OF INVENTION: Cellulase Variants  
FILE REFERENCE: 4887,204-US  
CURRENT APPLICATION NUMBER: US/09/261,329  
CURRENT FILING DATE: 1999-03-03  
EARLIER APPLICATION NUMBER: 1013/96  
EARLIER FILING DATE: 1996-09-17  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 202  
TYPE: PRT  
ORGANISM: cellulase variants  
US-09-261-329-3



Query Match 34.6%; Score 698; DB 9; Length 202;  
 Best Local Similarity 59.6%; Pred. No. 3.4e-40;  
 Matches 124; Conservative 30; Mismatches 46; Indels 8; Gaps 5;

QY 159 GAGVTRRYWDCCAGSWSFGKANVSSPVKSCNKGDTALSDSNVQSGC-NGNSIYCNAN 217  
 DB 1 GGRTRRYWDCCAGSWSFGKANVSSPVKSCNKGDTALSDSNVQSGC-NGNSIYCNAN 58  
 QY 218 OPAVAVNDNLAYGFAAALISGGGSRWCCSFELTFTSTVAGKMYIQTNTGGDLGSSST 277  
 DB 59 OPAVAVNDNLAYGFAAALISGGGSRWCCSFELTFTSTVAGKMYIQTNTGGDLGSSST 115  
 QY 278 GAHFDLQMPGGVGIENGSKOMGAPNDGSRVGISSASDSSLPALQAGCKRPMW 337  
 DB 116 GAHFDLQMPGGVGIENGSKOMGAPNDGSRVGISSASDSSLPALQAGCKRPMW 173  
 QY 338 FRKADNPSTYKEVTCPKEITAKTGCGR 365  
 DB 174 FNDADNPSTYKEVTCPKEITAKTGCGR 201

RESULT 14  
 US-08-841-636A-31  
 Sequence 31, Application US/08841636A  
 Patent No. US20020168751A1  
 GENERAL INFORMATION:  
 APPLICANT: Mettinen-Oinonen, Arja  
 APPLICANT: Londeeborough, John  
 APPLICANT: Vehmanen, Jari  
 APPLICANT: Haakana, Hei  
 APPLICANT: M ntyl, Arja  
 APPLICANT: Lantto, Raija  
 APPLICANT: Elvoinio, Minna  
 APPLICANT: Joutsinki, Vesa  
 APPLICANT: Paloheimo, Marja  
 APPLICANT: Suominen, Pirkko  
 TITLE OF INVENTION: NOVEL CELLULASES, THE GENES ENCODING THEM AND  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 STREET: 1100 New York Avenue, N.W., Suite 600  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/841,636A  
 FILING DATE: 30-APR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/005,335  
 FILING DATE: 17-OCT-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/007,926  
 FILING DATE: 04-DEC-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/020,840  
 FILING DATE: 28-JUN-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/732,181  
 FILING DATE: 16-OCT-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/FI96/00550  
 FILING DATE: 17-OCT-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Timothy J. Shea, Jr.  
 REGISTRATION NUMBER: 41,306

REFERENCE/DOCKET NUMBER: 1716.0510005/MAC/TJS  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)371-2600  
 TELEFAX: (202)371-2540  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 235 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: Melanocarpus albusmyces  
 STRAIN: ALKO4237  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..235  
 OTHER INFORMATION: /label= 20K-cellulase  
 US-08-841-636A-31

Query Match 34.2%; Score 690.5; DB 1; Length 235;  
 Best Local Similarity 56.6%; Pred. No. 1.3e-39;  
 Matches 120; Conservative 34; Mismatches 49; Indels 9; Gaps 4;

QY 155 GAGSNGVTRRYWDCCAGSWSFGKANVSSPVKSCNKGDTALSDSNVQSGCNGNSIYCN 214  
 DB 18 GALAANGSTRRYWDCCAGSWSFGKANVSSPVKSCNKGDTALSDSNVQSGCNGNSIYCN 76  
 QY 215 NDNQPAVNDNLAYGFAAALISGGGSRWCCSFELTFTSTVAGKMYIQTNTGGDLG 274  
 DB 77 ADHSPVAINDNLAYGFAAALISGGGSRWCCSFELTFTSTVAGKMYIQTNTGGDLG 136  
 QY 275 SNTGAHFDLQMPGGVGIENGSKOMGA-PNDGWSRVGISSASDSSLPALQAGCKR 333  
 DB 137 SN--HFDLNPQGGVGIENGSKOMGA-PNDGWSRVGISSASDSSLPALQAGCKR 189  
 QY 334 RPNWPKADNPSTYKEVTCPKEITAKTGCGR 365  
 DB 190 RPNWPKADNPSTYKEVTCPKEITAKTGCGR 221

RESULT 15  
 US-09-735-787-4  
 Sequence 4, Application US/09735787  
 Patent No. US20010036910A1  
 GENERAL INFORMATION:  
 APPLICANT: Raasmussen, Grethe  
 APPLICANT: Mikkelson, Jan Moller  
 APPLICANT: Schulten, Martin  
 APPLICANT: Paekar, Shankant A.  
 TITLE OF INVENTION: A Cellulase Preparation Comprising an  
 NUMBER OF SEQUENCES: 33  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. US20010036910A1 No. US20010036910A1disk of No. US2001003691  
 STREET: 405 Lexington Avenue, 64th Floor  
 CITY: New York  
 STATE: New York  
 COUNTRY: United States of America  
 ZIP: 10174-6401  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/735,787  
 FILING DATE: 13-Dec-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/189,028  
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Tamburris, Elias

NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33

REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 34

REFERENCE/DOCKET NUMBER: 3469.214-US  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 4:

### SEQUENCE CHARACTERISTICS:

LENGTH: 376 amino acids

TYPE: amino acid  
TOPOLOGY: linear

TOPOLOGY: linear  
CITE TYPE: protein

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION:

SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
15-787-4

US-09-135-181-4

Query Match	Score	DB	Length
33.8%	683.5	10	376

Best Local Similarity 57.1%; Pred. No. 6.2e-39;  
Matches 120; Conservative 34; Mismatches 49; Indels 7; Gaps 4;

```
Oy      157 ASGNGVTRRYMDCCKASCSWPGKANTSPSPVKSCNKGDTALISDSIVSGCGNGAGNS-YMCN 215
        |||:::||||| ||| ||| ::::: ||| |
Db      18 ASGGSHSTRYWDCCCKSCSWSGRKAANAPALTCDKND-NPISNTNAVNGCGGGSAVACT 76
```

016 DNOPWAVNDNI.AYGFAAAI SGGESRWCCSCEFLTFTSTSVACKMVI OVTNTGGDLGS 27

```

QY      216  DNQPMVNDNLAYGFAPAAAIISGGGSEKWCSCFELTITSTSVAGKRWIQLVNTGGDLS 275
      : ||||| ||||| ||||| : ||||| : ||||| : |||||
Db      77  NYSFPMVNDNLAYGFAPATKISGGSELSWCACVALTFTGPKVKKMIQSTNTGGDLD 136

```

376 STCAHEDI OMPCGCVITFNGCSKOWGAPNDGWGSPRYGGTSSASDCSSLPALOGCKWRF 335

```
QY      276 STGAHFDLPMGGGAVGI FNGCKKQMGA PNDGMSR YGGI SSASD CSSLP SALOAGCKKRF 335
       : ||||| |::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     137 N---HFDLMMPGGGVGIFDGCTSEFKALG--GAQYGGISRSRECDSYPELLKDGCWRF 191
```

336 NIEKVNADNDSCMTYKFEITCDEKITAKTGCSB 365

```
QY      336 NWFKADNPMSMTYKEVTCPEKITAKTGSGR   365  
       :|::||| |::| ||| : |||  
Db     192 DMFENADNPDFTEEQVOCPKALLDISGKR   221
```

Search completed: June 18, 2003, 17:44:42  
Job time : 80.5143 secs